Aging on RT-Level – Analysis and Monitoring

Ulf Schlichtmann

Institute for Electronic Design Automation
Technische Universität München (TUM)

ERMAVSS (Dresden, March 18, 2016)

Aging on (Gate and RT-Level) -

Analysis and Monitoring

Ulf Schlichtmann

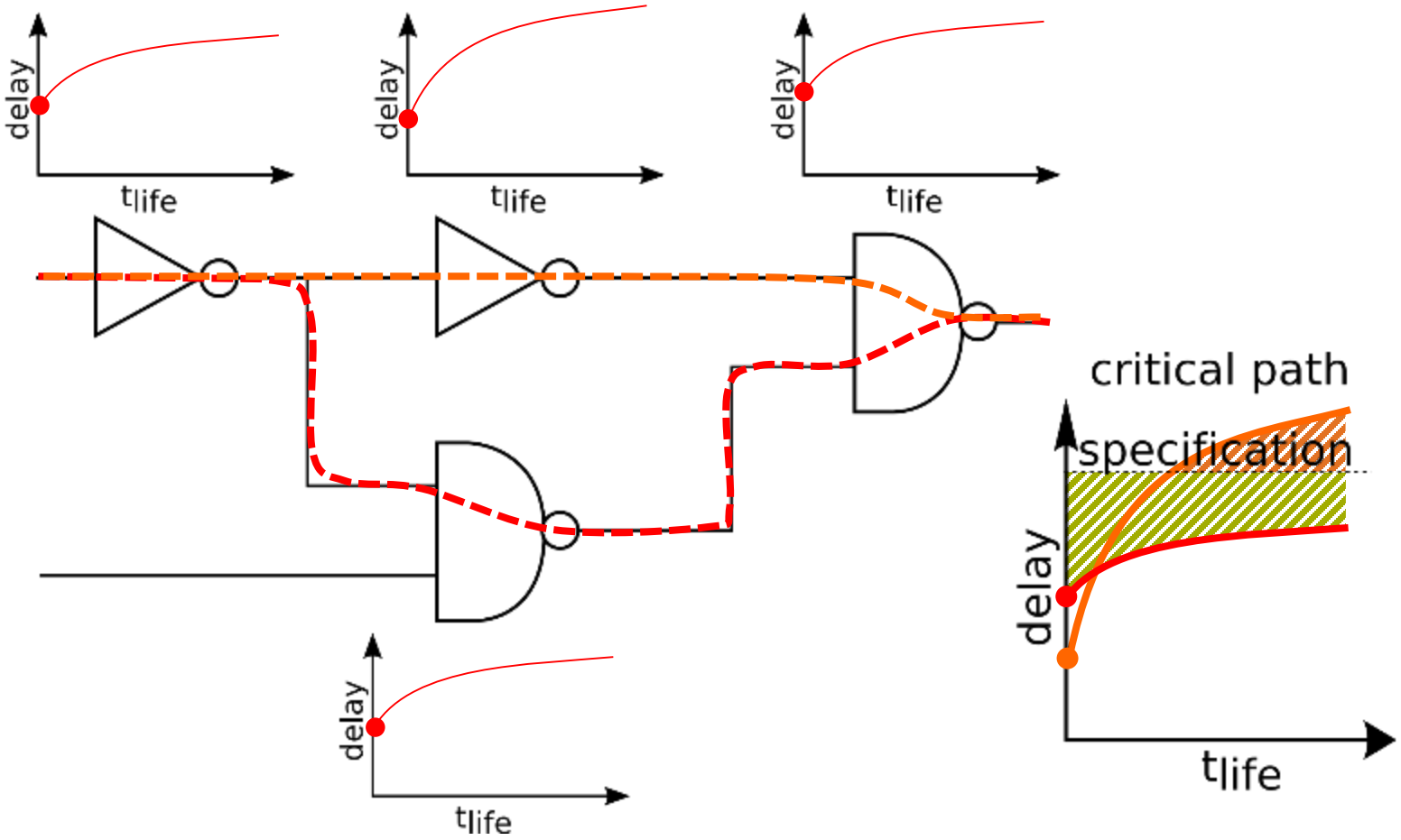
Institute for Electronic Design Automation
Technische Universität München (TUM)

ERMAVSS (Dresden, March 18, 2016)

Overview

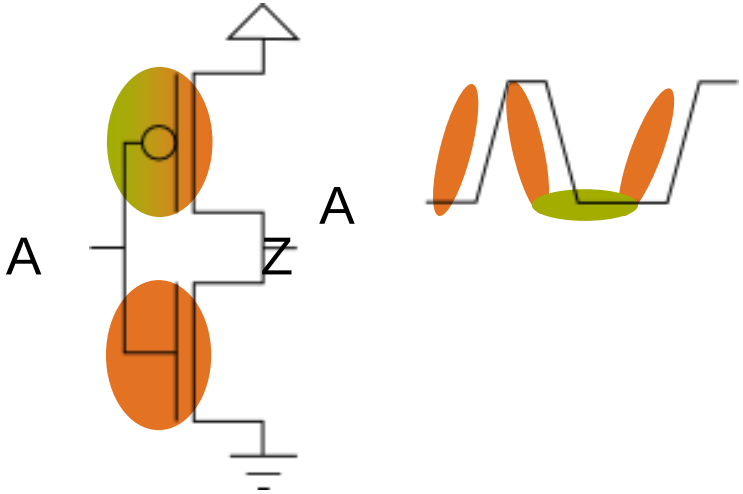
- Motivation
- Aging analysis on gate level
- Identifying possible critical paths
- Summary

(Aging-aware) timing analysis

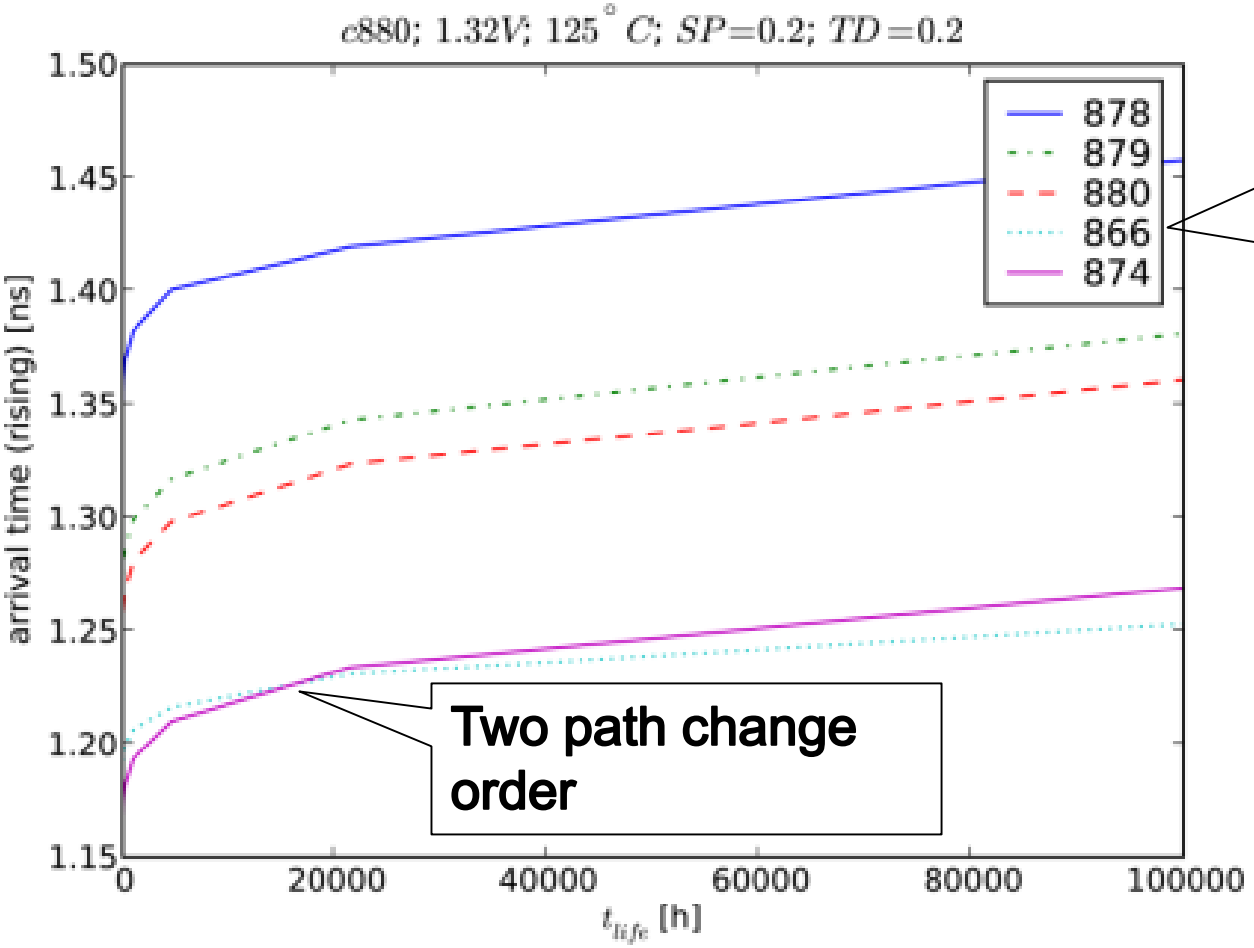


Dominant drift-related aging effects

	Negative Bias Temperature Instability (NBTI)	Hot Carrier Injection (HCI)
Device	PMOS	PMOS & NMOS
Modeled by	Threshold voltage drift ΔV_{th}	Degradation of drain saturation current ΔI_{on}
Stress condition	Transistor in inversion	Transistor switches



Critical path can change due to aging



Five slowest outputs of circuit c880

- Operating conditions: P_{nom}, 27° C, 0.9V
- Use profile: 125° C, 1.32V
- Workload by probabilistic method

Overview

Motivation

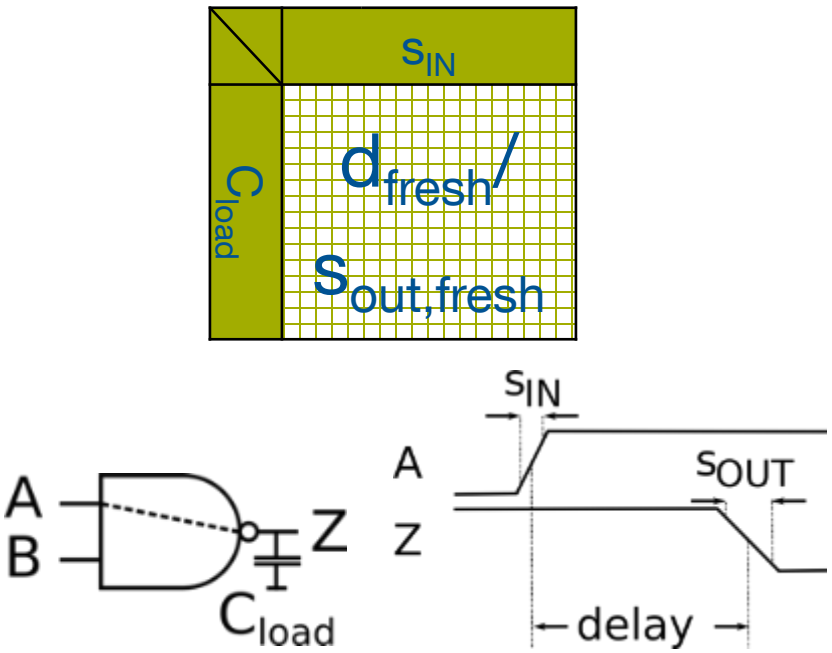
Aging analysis on gate level

Identifying possible critical paths

Summary

Aging analysis requires an aging-aware gate model

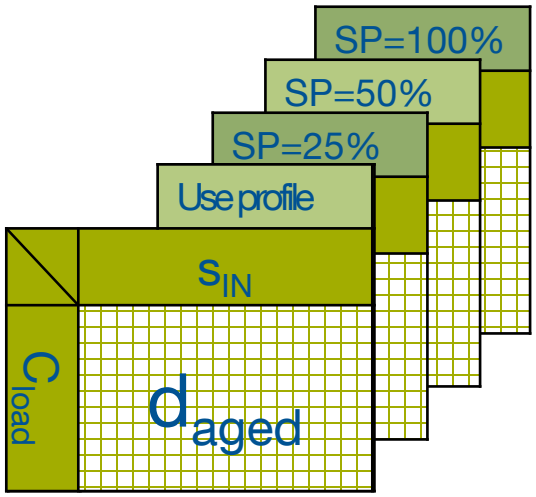
Traditional LUT-based gate model:



Additional dependencies for an aging-aware gate model:

- **Use profile**
  - Supply voltage over lifetime
  - Temperature over lifetime
- **Workload**
  - Signal probability (SP):
    Probability signal is logic 1
  - Transition density (TD):
    # of signal transitions

State of the art:
Aged look-up tables (Chen*)

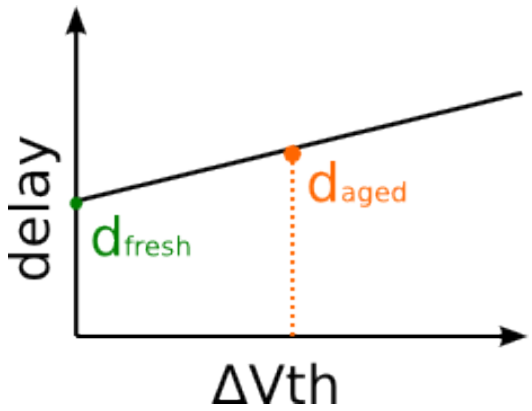


- ☺ Existing analysis flow can be reused
- ☹ New use profile ➔ library re-characterization
- ☹ Characterize gate for varying workload conditions

*“A Framework for Fast an Accurate Critical-Reliability Paths Identification”, Chen et al., North Atlantic Test Workshop 2011.

State of the art:
Aged gate delay as function of parameter drift (Paul*)

$$d_{aged} = d_{fresh} + \frac{\partial d}{\partial V_{th}} \cdot \Delta V_{th}$$



- ☺ Use profile independent

- ☹ Only delay, no slope
- ☹ Only threshold voltage drift ΔV_{th}
- ☹ One ΔV_{th} for all transistors

- Sensitivity by α-power-law model:

$$\frac{\partial d}{\partial V_{th}} = \frac{\alpha}{V_{gs} - V_{th}} \cdot d_{fresh}$$

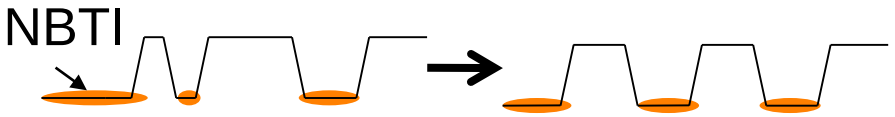
*"*Temporal Performance Degradation under NBTI: Estimation and Design for Improved Reliability of Nanoscale Circuits", Paul et al., DATE'06*

State of the art:
Aged gate delay as function of parameter drift (cont.)

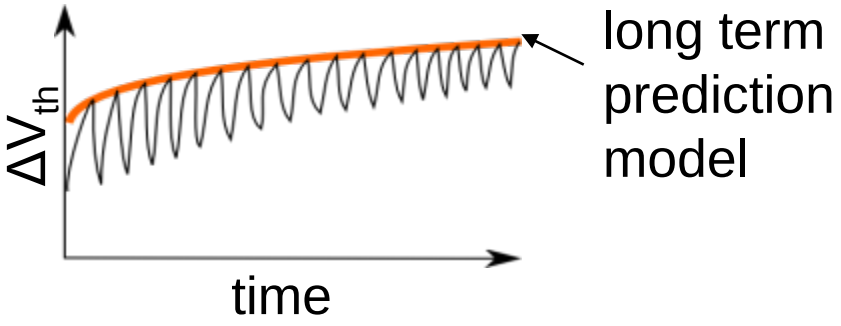
- Dependence on ΔV_{th} obtained during characterization (Kumar*):

$$d_{aged} = d_{fresh} + \Delta d(\Delta V_{th})$$

- ΔV_{th} depends on SP not on actual waveform (Kumar*)

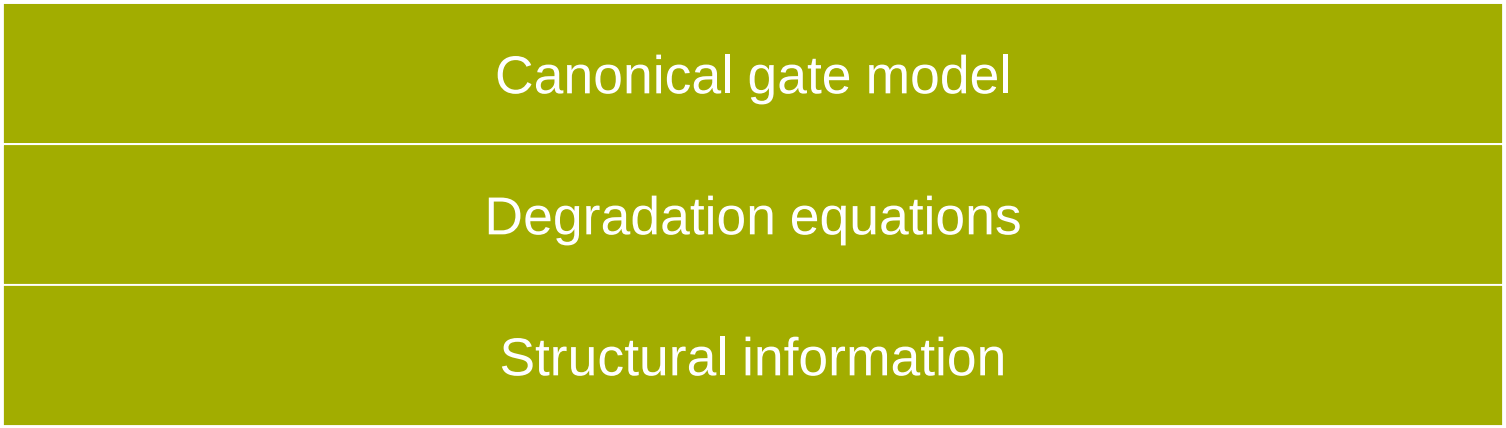


- Long term prediction model for ΔV_{th} : closed form for upper bound of reaction-diffusion model (Wang**)



*"An Analytical Model for Negative Bias Temperature Instability", Kumar et al., ICCAD'06
**"The impact of NBTI on the Performance of Combinational and Sequential Circuits", Wang et al., DAC'07

Aging-aware gate model: AgeGate*



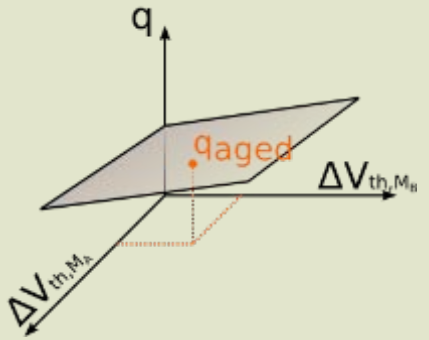
*"Aging Analysis of Circuit Timing Considering NBTI and HCI", Lorenz et al., IEEE International On-Line Testing Symposium (IOLTS) 2009

Gate model provides aged gate performance for drifts

Canonical gate model

$$q_{aged} = q_{fresh} + \sum_{\substack{m \in \\ \text{transistors} \\ \text{in gate}}} \chi^{q}_{\Delta V_{th},m} \cdot \Delta V_{th,m} + \chi^{q}_{\Delta I_{on},m} \cdot \Delta I_{on,m}$$

q_{aged}, q_{fresh}: aged, fresh gate performance
χ: sensitivity coefficient



Degradation equations

Structural information

Drifts calculated by degradation equations

Canonical gate model

Degradation equations

- NBTI: $\Delta V_{th} = f(UP, t_{stress}, W, L)$
- HCI: $\Delta I_{on} = f(UP, t_{stress}, W, L)$

- $t_{stress} = P_{stress} \bullet t_{life}$

Workload

Internal gate structure

UP: use profile

P_{stress}: stress probability

W,L: transistor sizes

Structural Information

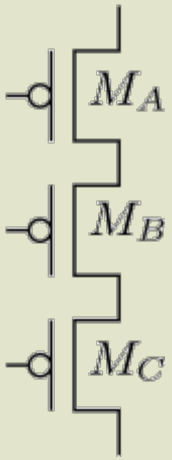
Structural information needed for drift calculation

Canonical gate model

Degradation equations

Structural information

- Transistor sizes
- Transistor stack structure



Calculating stress probability for NBTI

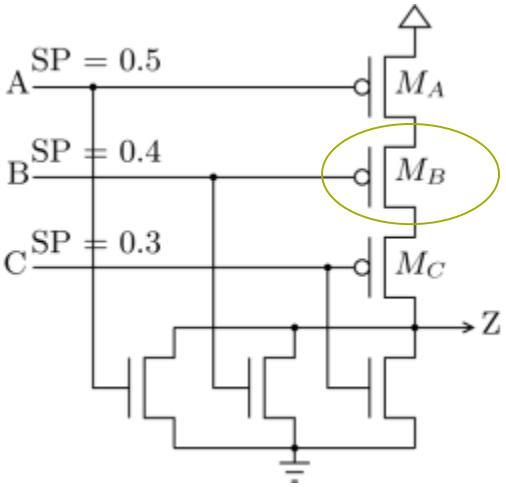
Stress condition for transistor M:

- C_1: logic “0” applied to gate terminal of M
- C_2: logic “1” applied to source or drain terminal of M

$$P(C_1) = 1 - SP(M)$$

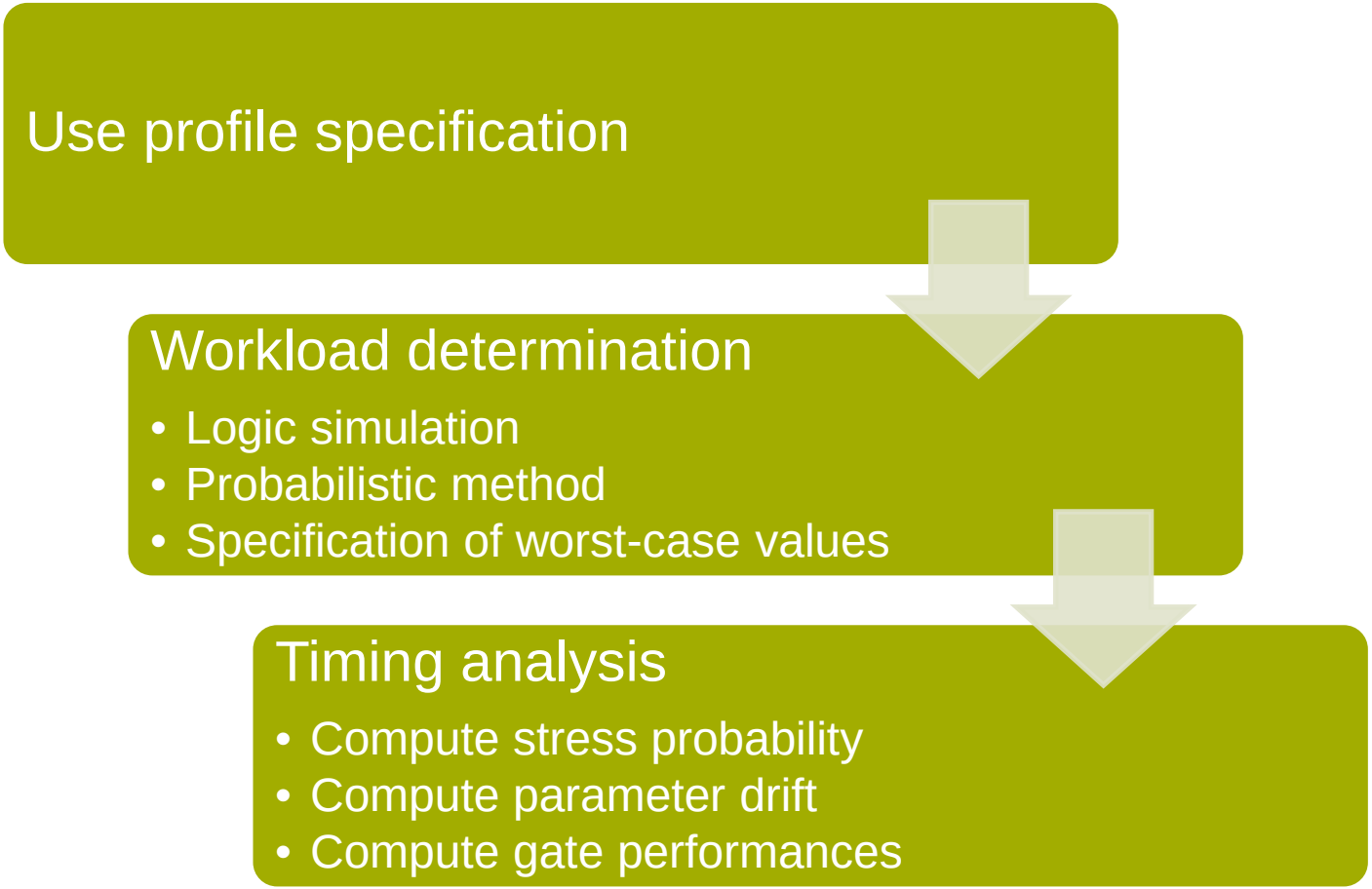
$$P(C_2) = \prod_{t \in PATH_{NBTI}} 1 - SP(t)$$

NOR3



$$P_{\text{stress},NBTI} = P(C_1 \wedge C_2) = P(C_1) \cdot P(C_2)$$

Aging analysis flow

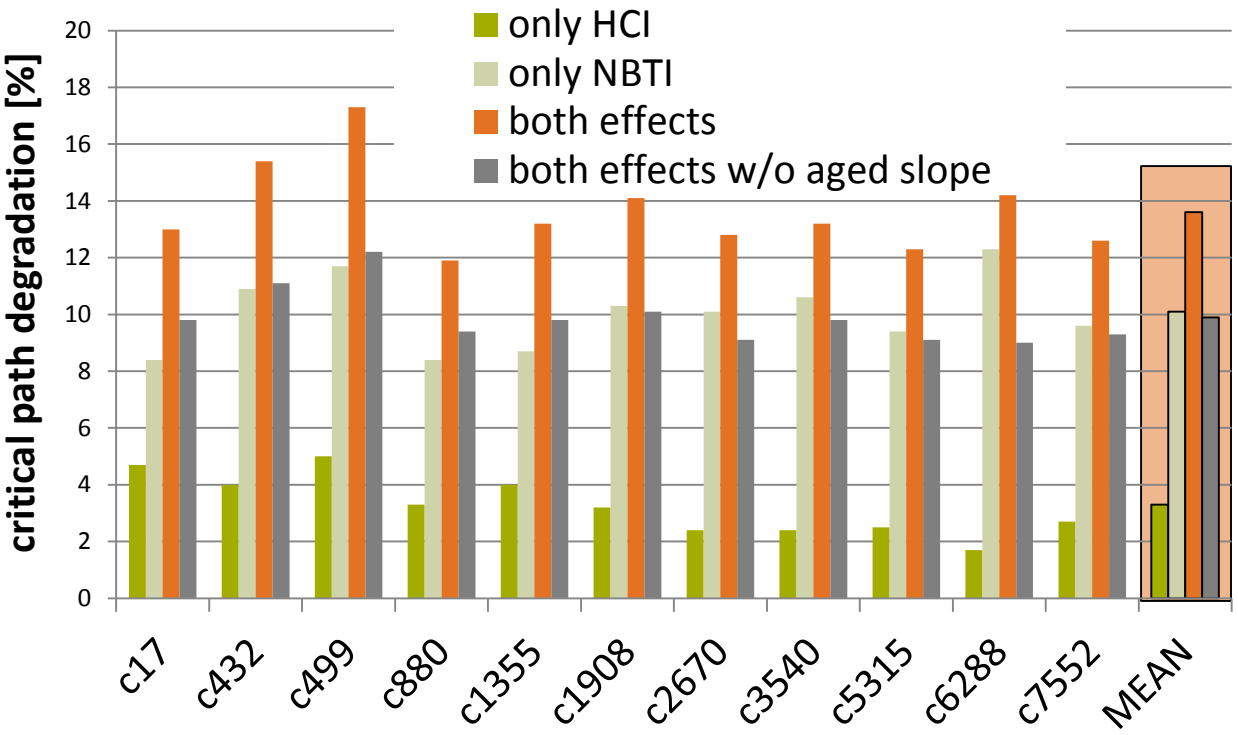


*"Efficienly analyzing the impact of aging effects on large integrated circuits", Lorenz et al., Microelectronics Reliability 52(8), 1546-1552

AgeGate compared to state-of-the-art

	Chen‘11	Paul‘06	Kumar‘06	Wang‘07	AgeGate
NBTI	✓	✓	✓	✓	✓
HCI	✓				✓
Individual transistor drifts	✓				✓
Aged output slope	✓				✓
Use profile independent model		✓	✓	✓	✓

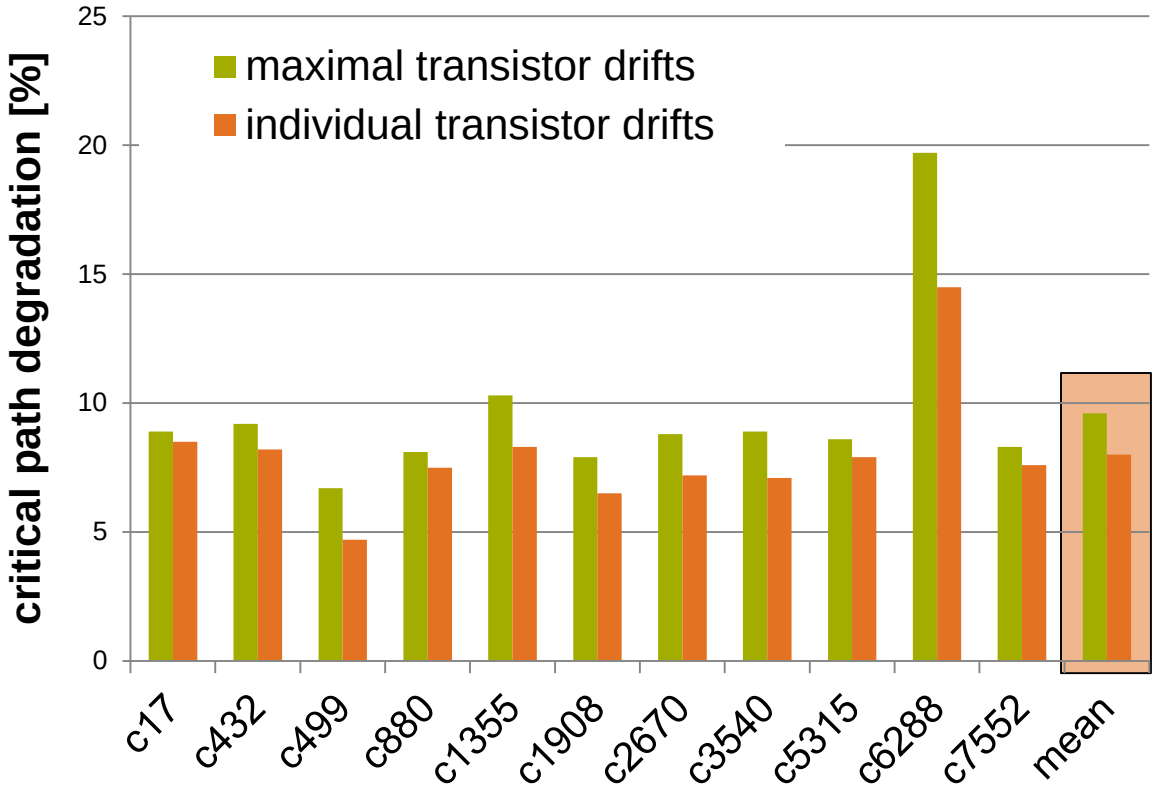
Degradation of critical path delay for ISCAS'85 benchmark circuits



- Industrial 90nm cell library
- Use profile: 125° C, 1.32V, 10y
- Measurement conditions: 27° C, 0.9V
- Worst-case analysis (SP=0 and TD=2 for all nets)
- Runtime: 35s for c7552

- Both effects are relevant
- Not considering aged output slope ➔ 24% underestimation

Comparison: w and w/o individual transistor drifts



- Workload estimation with probabilistic method (*SP*=0.5 and *TD*=0.4 at all inputs)

- w/o individual transistor drifts ➔ degradation 20% overestimated

Summary – Gate Level

- Accurate aging analysis methods to reduce safety margins

- Accuracy of aging-aware gate model depends on
  - Considering all dominant aging effects
  - Individual transistor drifts
  - Aged output slope

Overview

Motivation

Aging analysis on gate level

Identifying possible critical paths

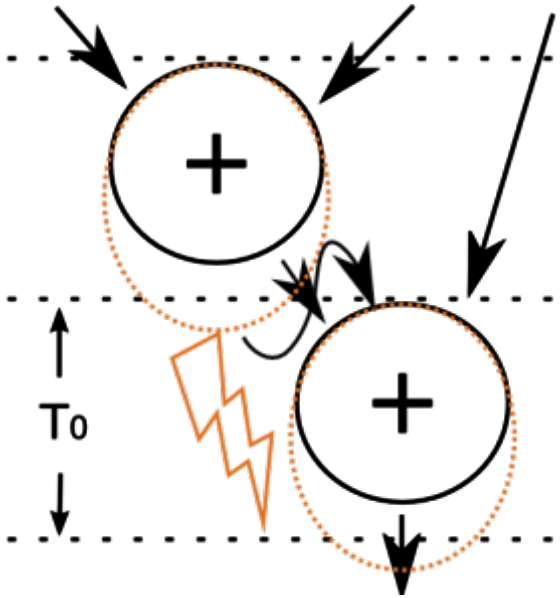
Summary

Without knowing operating conditions or workload, a definite aging-analysis is not possible

- Possible critical paths (PCPs):
  Path that might become critical during the lifetime

- Applications:
  - Timing model on system-level
  - Monitoring of aging circuits

Application of PCPs: Design Space Exploration

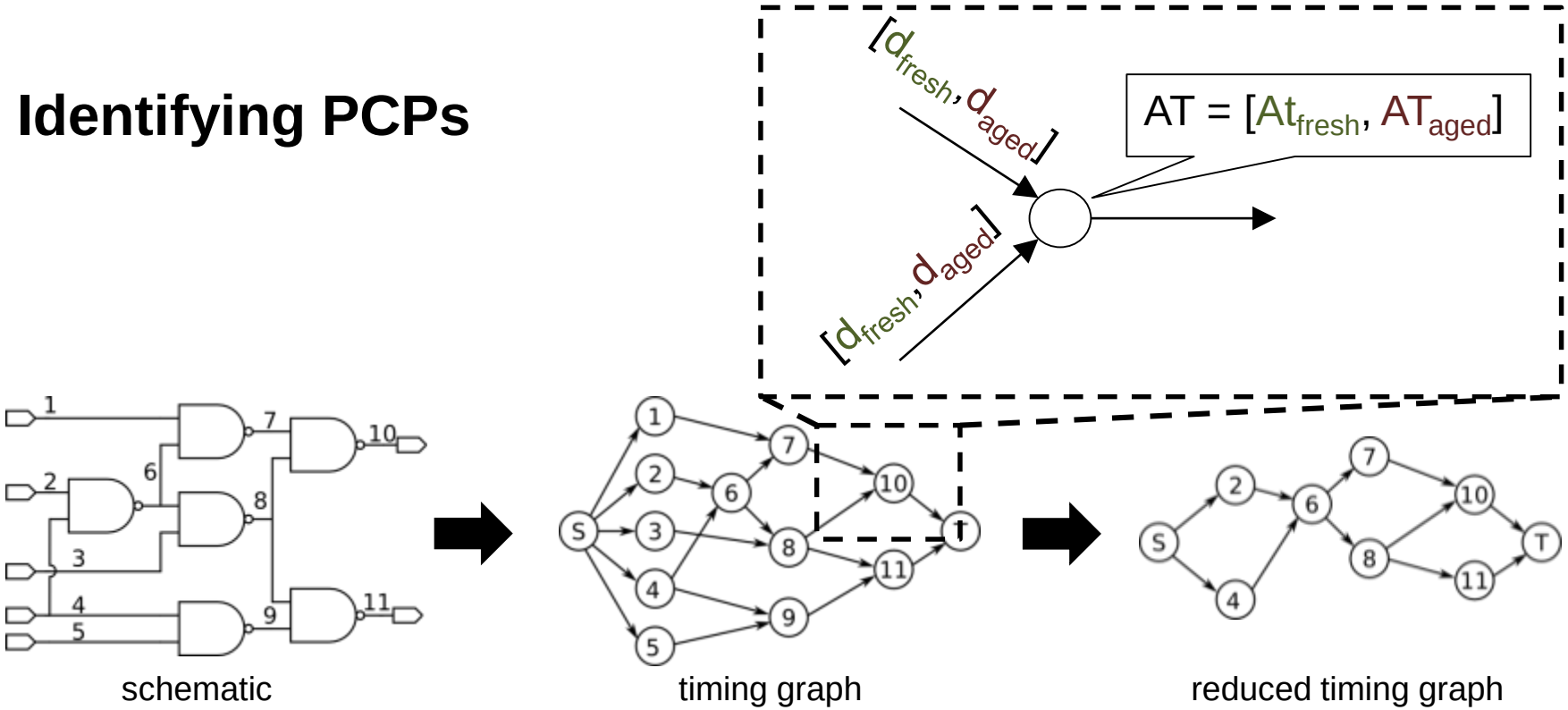


- Scheduling: Mapping of logical / arithmetic operations to time slot of duration T_0
- Aged circuit fails because operation takes longer than T_0

Aging-aware timing model (TM) at RTL enables:

- Considering impact of aging on system early in design process
- Quick performance determination at system level
- Design space exploration

Identifying PCPs

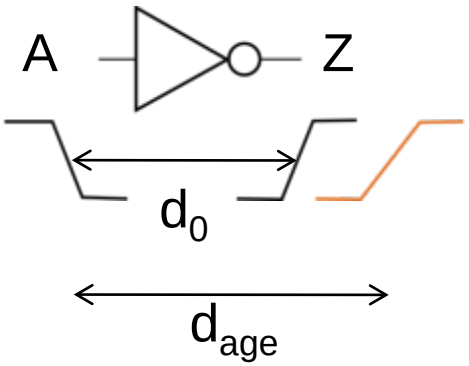


Basic idea: Reduced timing graph (TG)

- Gate delays modeled as intervals
- Identify and remove nodes and edges not part of a PCP

Aging-aware gate model needed for generation/evaluation of TM

operating cond. over lifetime (use profile) & workload

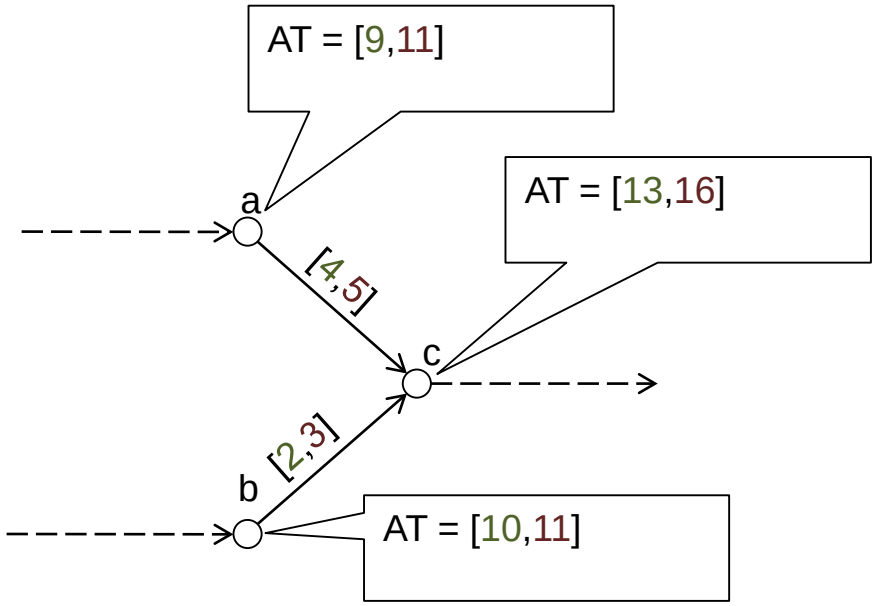


- Approach independent of used gate model
- Interval for gate delay because actual gate delay unknown during characterization

d_0: fresh gate delay

d_{age}: aged gate delay

Reducing the TG

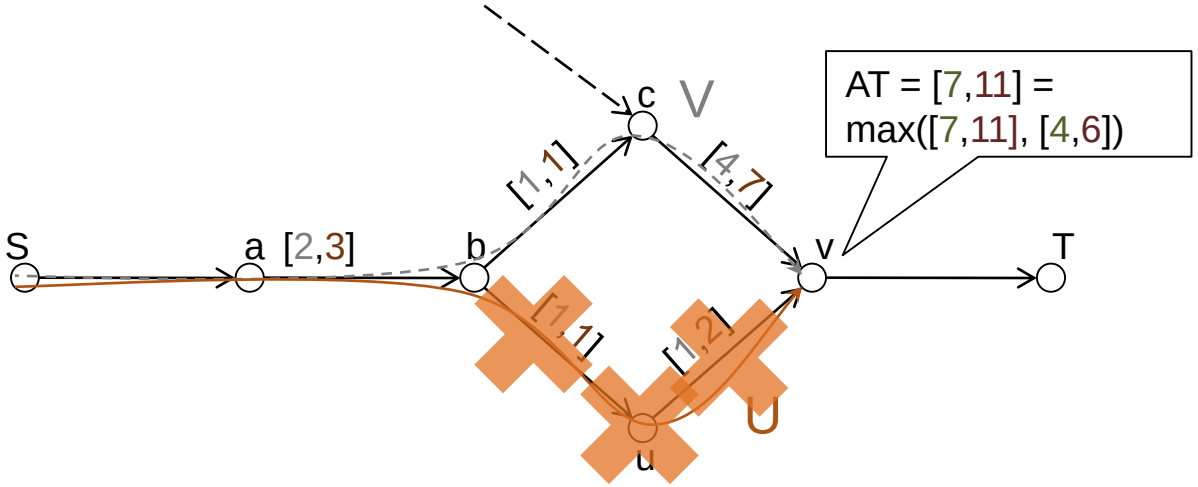


Annotated timing graph

- Identify elements not part of a possible critical path:
  1. Slack reduction step
  2. Path delay reduction step
  3. Arrival time/Delay to sink reduction step
  4. Common edge reduction step
  5. Path-based reduction step

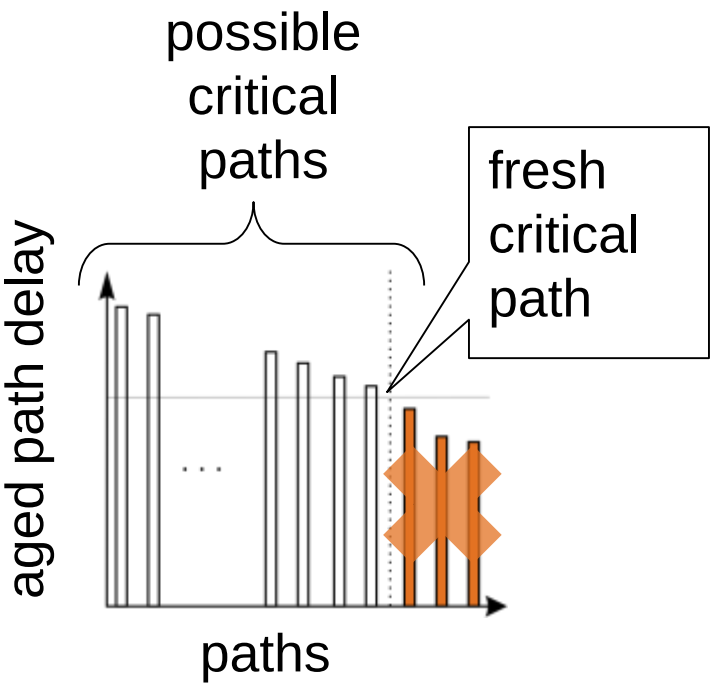
**"Efficienly analyzing the impact of aging effects on large integrated circuits", Lorenz et al., Microelectronics Reliability 52(8), 1546-1552*

Arrival time reduction step



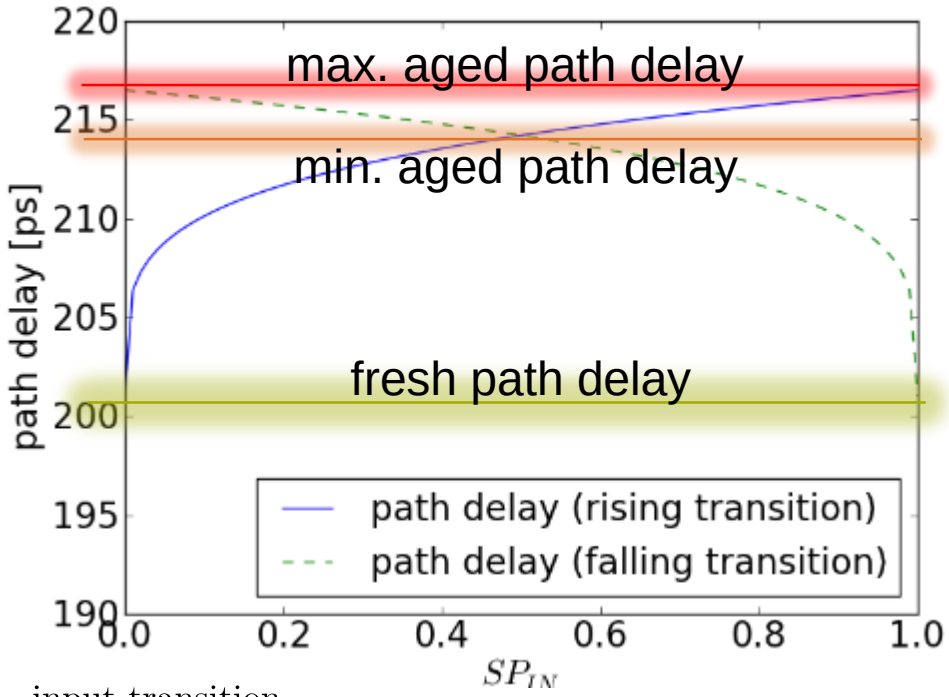
- Static timing analysis with intervals

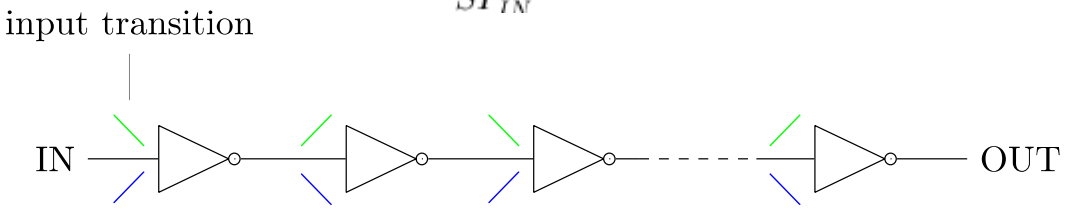
Path-based reduction step



- Aged delay of PCP > fresh delay of nominal critical path!
- Enumerate all paths with aged path delay greater than $D(P_{crit})_{fresh}$

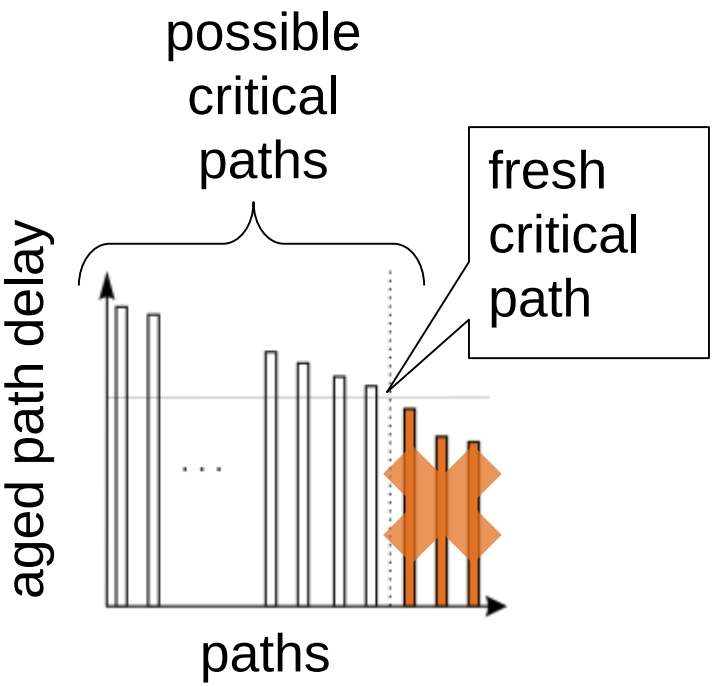
Minimal aged path delay





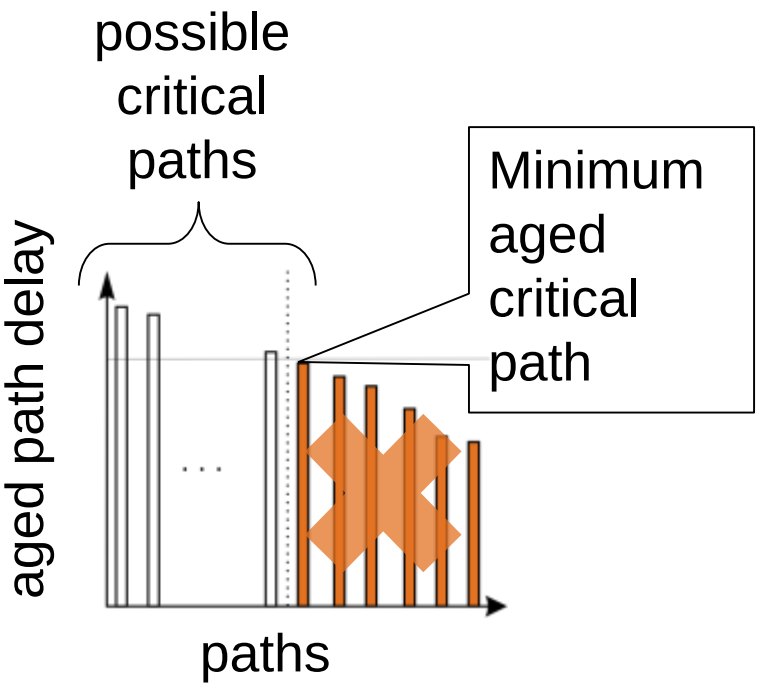
- (Most) circuits degrade independent of the actual workload
- Simple example: inverter chain
- Linear optimization problem formulated to obtain lower limit for minimal aged circuit delay

Path-based reduction step with minimal aged circuit delay



- Minimum aged circuit delay used instead of fresh circuit delay

Path-based reduction step with minimal aged circuit delay



- Minimum aged circuit delay used instead of fresh circuit delay

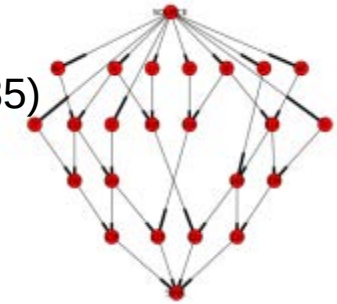
Results

- RC-adder, CLA-adder, ISCAS'85 benchmarks
- 90 nm industrial cell library

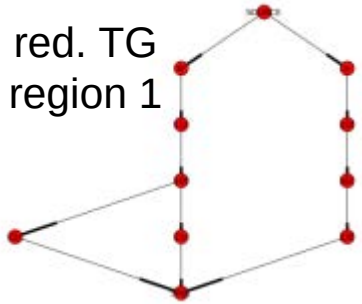
- $$\text{Reduction (ratio)} = \frac{\text{removed nodes (edges)}}{\text{nodes (edges) of original TG}}$$

- Two validity regions specified:
  - Region 1: 100° C; 1.2V; 5y ➔ max ΔV_{th} = 8%
  - Region 2: 125° C; 1.32V; 10y ➔ max ΔV_{th} = 17%

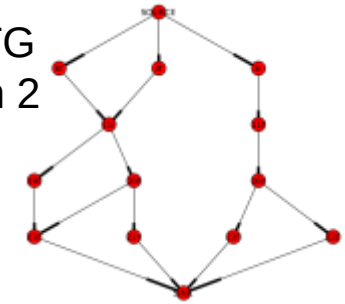
original TG
C17(ISCAS'85)



red. TG
region 1



red. TG
region 2



PCPs could be reduced by drastically

Circuit	# Gates	# Paths	# PCPs State-of-the-art *	# PCPs This approach	Reduction
c499	534	452608	1487	375	4.0×
c1355	589	522368	3376	2224	1.5×
c2670	708	31286	21	21	1.0×
c3540	905	4248254	15276	1345	11.4×
c5315	1484	738816	1568	899	1.7×
c6288	2601	5.1e+16	6.8e+12	4.1e+12	1.6×
c7552	2242	448564	3173	522	6.1×

Mean: 4×

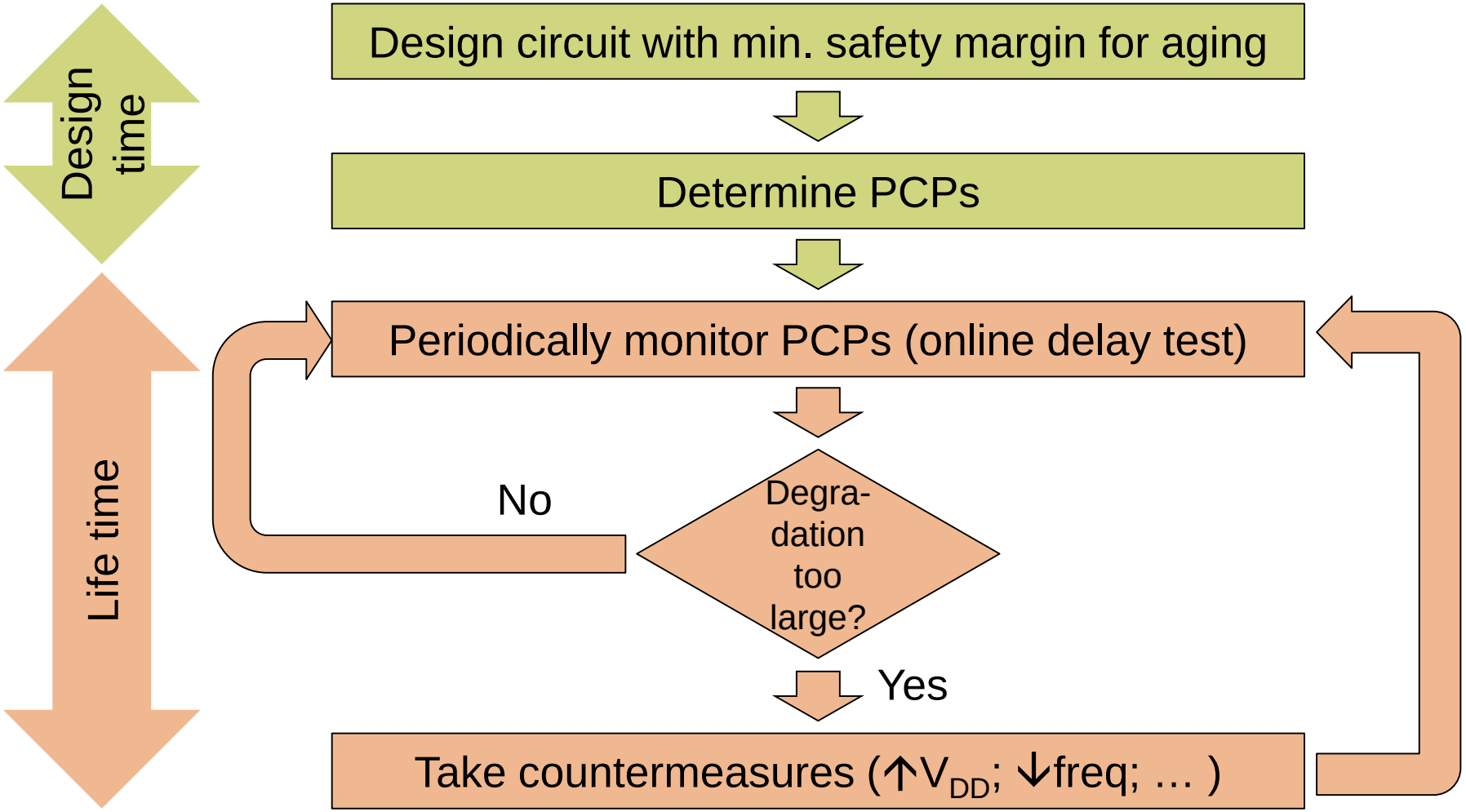
- 90 nm cell library
- 125° C; 1.32V; 10y
- Only NBTI
- Circuits with # Gates > 500

* “Testing for transistor aging”, Baba et al., VTS’09

Ways to monitor aging

- Single device monitors
  - Measure threshold voltage drift
  - ☹ Hard to correlate to circuit performance
- Generic test structure (e.g. inverter ring oscillator)
  - Measure oscillating frequency
  - ☹ Neglects workload impact
- Circuit replica (e.g. critical path replica)
  - Measure delay/oscillating frequency
  - ☹ Neglects workload impact
- On-line delay fault testing
  - Measure delay of current critical path
  - ☺ Only safe way
  - ☹ Circuit operation must be paused

Application of PCPs: Monitoring aging circuits



*"Monitoring of Aging in Integrated Circuits by Identifying Possible Critical Paths", Lorenz et al., Microelectronics Reliability 54(6-7), 1075-1082

Ways the system can react

- Disable degraded circuit (e.g. one core of multi-core processor)
- Reduce clock frequency
- Increase supply voltage (Caution: accelerates aging)
- Replace degraded circuit by redundant one (degraded circuit can recover)
- Use degraded circuit for uncritical tasks (probabilistic CMOS)

Overview

Motivation

Aging analysis on gate level

Identifying possible critical paths

Summary

Summary

- Accurate aging analysis methods to reduce safety margins
- Accuracy of aging-aware gate model depends on
  - Considering all dominant aging effects
  - Individual transistor drifts
  - Aged output slope
- Identify possible critical paths to monitor circuit degradations
  - No large safety margins, instead react if degradation gets too large
  - Number of PCPs reduced by 4× compared to state-of-the-art approach

TUM. Technische Universität München